

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 05:58:15 ; Search time 2251.99 Seconds
(Without alignment)
12849.734 Million cell updates/sec

Title: US-09-821-839-1
Perfect score: 2144
Sequence: 1 acgcacacagccacactctc.....aaaaaaaaaaaaa 2144

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: em_estba:*
2: em_estbha:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_hic:*
13: em_gss_hum:*
14: em_gss_hum:*
15: em_gss_hum:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	609.8	28.4	1233	12 B10183	B10183 F19N12-SP6
2	188.8	8.8	192	12 B97951	B97951 F22C23TFB I
3	188.8	8.8	818	12 BH530370	BH530370 BOGDO70TF
4	175	8.2	531	10 BF291837	BF291837 WHE2204_E
5	172.6	8.1	1166	12 B12531	B12531 F1C14-T7 IG
6	165.2	7.7	803	12 BH530380	BH530380 BOGDO70TF
7	152.4	7.1	465	12 BH462012	BH462012 BOHNA23TF
8	147.4	6.9	727	12 BH487741	BH487741 BOGMB48TF
9	116.8	5.4	336	12 CANSO04M	BM270448 Arabidops
10	96.4	4.5	498	10 BM270448	BM270448 sak13a03
11	93.2	4.3	612	12 BH477964	BH477964 BOHNA23TF
12	80.8	3.8	605	12 BH462017	BH462017 BOHNA23TF
13	70.6	3.3	791	12 A2686821	A2686821 ENTU174TR
14	70.4	3.3	668	9 AL514901	AL514901 Drosophila
15	68.8	3.2	1225	12 CANSO16ID	BE420736 HMM002.A0
16	68.4	3.2	1309	10 BE420736	AL060732 Drosophila
17	68	3.2	1101	12 CANSO01FB	

C 18	67.8	3.2	330	9 AL513817	AL513817
C 19	67.8	3.2	634	9 AL514047	AL514047
C 20	67.6	3.2	410	9 AL513713	AL513713
C 21	67.6	3.2	590	9 AL515321	AL515321
C 22	67.4	3.1	423	9 AM119922	AM119922 sds4521
C 23	67.2	3.1	613	12 A0922590	A0922590 RPCI-23-2
C 24	67	3.1	661	12 CANSO40UH	AL303074 Tetraodon
C 25	66.2	3.1	987	12 CANSO14PO	AL104456 Drosophila
C 26	65.8	3.1	388	12 CANSO39KI	AL234027 Tetraodon
C 27	65.8	3.1	401	9 AL515191	AL515191
C 28	65.6	3.1	329	9 AL513719	AL513719
C 29	65.6	3.1	456	9 AL513913	AL513913
C 30	65.6	3.1	523	9 AL514015	AL514015
C 31	65.6	3.1	619	9 AL514935	AL514935
C 32	65.4	3.1	329	9 AL513719	AL513719
C 33	65.4	3.1	421	10 BM270739	BM270739 sak17a05
C 34	65.2	3.0	714	12 A0324694	A0324694 mxyb0019D
C 35	65.2	3.0	777	12 CANSO25WB	AL182612 Tetraodon
C 36	65.2	3.0	988	12 CANSO072R	AL066743 Drosophila
C 37	65	3.0	250	10 B1742555	B1742555 Kt52d10.Y
C 38	64.6	3.0	230	9 A1811812	A1811812 tw32d10.X
C 39	64.2	3.0	453	10 BP298095	BP298095 060PB12
C 40	64.2	3.0	618	10 BE674847	BE674847 601488792
C 41	64.2	3.0	767	12 CANSO04OX	AL055924 Drosophila
C 42	64.2	3.0	951	12 A2672893	A2672893 ENTU174TR
C 43	64	3.0	954	12 BHL62327	BHL62327 ENTRJ37TR
C 44	63.8	3.0	279	9 A1540752	A1540752 tn90d04.x
C 45	63.8	3.0	633	9 AL513979	AL513979

ALIGNMENTS

RESULT 1
LOCUS B10183/c
DEFINITION F19N12-SP6 IGF Arabidopsis thaliana genomic clone F19N12, DNA
ACCESION B10183
VERSION B10183.1 GI:2091302
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1233)
REFERENCE
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shim, P., Sun, H. and
Ecker, J.
BAC End Sequences at ATGC
Unpublished (1997)
Other GSSs: F19N12-T7
Contact: Ecker J.
Arabidopsis Thaliana genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 87
High quality sequence stop: 806.
Location/Qualifiers
1..1233
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F19N12"
/clone_lib="IGS"
/sex="thermaphrodite"
/note="vector: BclBACII; Site_1: EcoRI; Site_2: EcoRI;

FEATURES

source

OY	1550	tgcttcaccaccacaacatcctcgtctgtgtctctcctcaaaagctgcgcagccaat	1609
Db	184	tgctttgcctgacacagtcacacattttctctctgcttctctgtaaacgcacaaagcagac	243
OY	1610	cagagattgaaagagaaagcaaacctctgtgcgtgttaactcaactatccgacaaactcga	1669
Db	244	gacaaatgtagagatctgcgaaagcaccgcgcctgcctcacacattcttgagacattaaagac	303
OY	1670	ctctggtttttggccctccaactctgtagaagctctgaactcgtgtgtctcgcctgcatcgaaac	1729
Db	304	ctctctctacttgccctctgcacccgtgcgaagcctctggatggagcccttgcctctgtggccacc	363
OY	1730	aacaaatcttcctgcataccacaagagatcataaagatgcaltgttagaacacagabaacgag	1789
Db	364	gacacagagagctctcgtgcaccatgcagaagaaaggagactcctcgaaagacgaagagacagac	423
OY	1790	ttgcctgaaatgcgtlaagagctctgagactgtgtct	1824
Db	424	ctgcccgaatggtttaatgagctctggagctggcgcgt	458

RESULT	5
LOCUS	B12531
DEFINITION	B12531 1166 bp DNA linear GSS 14-MAY-1997
ACCESSION	FLC14-T7 1CF Arabidopsis thaliana genomic clone FLC14, DNA sequence.
VERSION	B12531
KEYWORDS	B12531.1 GI:2093651
SOURCE	GSS
ORGANISM	thale cress.
	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 1166)
REFERENCE	Feng, J., Dwarak, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
AUTHORS	
TITLE	BAC End Sequences at ATGC
JOURNAL	Unpublished (1997)
COMMENT	other_GSSs: FLC14-Sp6

Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@at.genome.bio.upenn.edu
Seq primer: 17
Class: BAC ends
High quality sequence start: 37
High quality sequence stop: 889.
Location/Qualifiers
1. 1166

	/organism="Arabidopsis thaliana"	
	/strain="Columbia"	
	/db_xref="taxon:3702"	
	/clone="FIC14"	
	/clone_1kb="IGF"	
	/sex="hermaphrodite"	
	/note="Vector: BclBACII; Site_1: EcoRI; Site_2: EcoRI"	
BASE COUNT	Produced by Thomas Altmann	
ORIGIN	247 c 234 g 375 t	
	308 a	2 others

```

Query Match      8 1%; Score 172.6; DB 12; Length 1166;
Best Local Similarity 86.8%; Pred. No. 3e-14;
Matches 190; Conservative 29; Indels 0; Gaps 0;
QY 1081 ttgtgacctctctcgaggaagaattcactcgaattgttaagctttatgacgagga 1140
      ||||||| | || | || | |||||

```

Db	297	TTGTGGAAATGAATCTTAAGTATACATTTTCGTTTAACTGCTCAAGCTTGATGATGATGAGA	356
Qy	1141	ggctggaagaagactatctctaagcctgagggaaagaaagaagactacgcatatctgcggga	1200
Db	357	GGTGGAAAGAGAGCTATCTTAAGGCTGAGGGGAAAGAAAGAAAGTCAATCATATATTCGGGGA	416
Qy	1201	ctgggctaaagccatacctgcgtccagagatggacaatactcgtgtctcatccctcgtctaacgctc	1260
Db	417	CTGTGCTTAAGGCATATACGCTCTCCAGAGATGGCAATACGTCGTCACTCCGCTGTACGCTC	476
Qy	1261	catcatggttcacatgagatctgtaaaagcaatgttctgcaat	1299
Db	477	CATCATGCTTCATATGATGTGTAAAGGTAATTTTAACTT	515

RESULT	6
BH530380/c	
LOCUS	
DEFINITION	BH530380 803 bp DNA linear GSS 14-DEC-2001 BOGDQ70TF BOGD Brassica oleracea genomic clone BOGDQ70, DNA
ACCESSION	
VERSION	BH530380
KEYWORDS	sequence.
SOURCE	BH530380.1 GI:17746944
ORGANISM	GSS. Brassica oleracea. Brassica oleracea
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 803)
AUTHORS	Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE	Whole genome shotgun sequencing of Brassica oleracea
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSs: BOGDQ70TF

Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 7R
Class: sheared ends.

FEATURES	Location/Qualifiers
source	1..803
	/organism="Brassica oleracea"
	/strain="TO1000DH3"
	/db_xref="taxon:3712"
	/clone="BOCD070"
	/clone_lib="BOCD"
	/note="Vector: Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS1 using BstXI linkers"
BASE COUNT	241 a 128 c 181 g 253 t
ORIGIN	

Query Match	7.78;	Score 165.2;	DB 12;	Length 803;
Best Local Similarity	69.08;	Pred. No. 3.3e-13;		
Matches 287; Conservative	0;	Mismatches 43;	Indels 86;	Gaps 1;

Oy 1451 agcattccggaanaagaaacttcaccatctcagaacctatagaccggcatgaagtgtg 1510
|||||
Db 614 AGCATCCGGAANAAGAACTTCTACTTTGAGAACCTTAAGTGTATGACCGTCATGATGAGTG 555

QY 1511 gcaatgagctgctggttcaagaagtcctcaactlcaaatgcttcacacccacaatcttc 1570

Db	554	GCAATGGAGTGGCTGATTTCTAGAGTCCCTCAACTTCAATGCTTCTCACCCACAAATCTTT	495
----	-----	---	-----

494	AACTCTTAAWGTAACCCCTATTACTATATATTTCTGTTCGAGACACATGCACAC	435
495		
496		
497		
498		
499		
500		
501		
502		
503		
504		
505		
506		
507		
508		
509		
510		
511		
512		
513		
514		
515		
516		
517		
518		
519		
520		
521		
522		
523		
524		
525		
526		
527		
528		
529		
530		
531		
532		
533		
534		
535		
536		
537		
538		
539		
540		
541		
542		
543		
544		
545		
546		
547		
548		
549		
550		
551		
552		
553		
554		
555		
556		
557		
558		
559		
560		
561		
562		
563		
564		
565		
566		
567		
568		
569		
570		
571		
572		
573		
574		
575		
576		
577		
578		
579		
580		
581		
582		
583		
584		
585		
586		
587		
588		
589		
590		
591		
592		
593		
594		
595		
596		
597		
598		
599		
600		
601		
602		
603		
604		
605		
606		
607		
608		
609		
610		
611		
612		
613		
614		
615		
616		
617		
618		
619		
620		
621		
622		
623		
624		
625		
626		
627		
628		
629		

QY 1581 -----gattctacttaagaagctgctgag 1604

```

Db          78  AGTAC  73

RESULT      8
 accession  BH487741
 LOCUS      BH487741
 DEFINITION sequence;
            BH487741
            BH487741.1  GI:17695845
            GSS.
 ORGANISM   Brassica oleracea.
            Brassica oleracea.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
            1 (bases 1 to 727)
 REFERENCE 1  Town, C.D., Van Aken, S., Ullrich, T., and Fraser, C.M.
            Whole genome shotgun sequencing of Brassica oleracea
            unpublished (2001)
            Other_GSSs: BOCMB487F
 COMMENT    TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtowne@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
            Location/Qualifiers
                1..727
                /organism="Brassica oleracea"
                /db_xref="taxon:10000"

```

BASE COUNT	188 a	168 c	196 g	175 t	ORIGIN
Query Match					genomic DNA inserted into pRO1 using Bsal linkers-
Best Local Similarity	67.9%				Score 147.4; DB 12; Length 727;
Matches 324; Conservative					Pred. No. 8,1e-11;
					Mismatches 66; Indels 87; Gaps 5
QY 639	cggagaaataagagagacgcagctcgttctcgtatatacgcggattgagacagttcccaagt				638
DB 233	CGGAGAAATAMAGGACAGCGACCTCG-----TATCGGAGATTAAGCTTGCTGCCTCAAGT				263
QY 639	tgggagcggttacccggggagagct-----gataacgaaagaattgaattcccaagc				749

QY	750	CGCAGCCTCTGGAGAGCGTATCCTCCTCTTGATCGCCAGGAATLGGACGCGAGC	809
Db	344	CGAGCGGACACAGTGGAGC-----GAAGCGGGAGT	373
QY	810	ttgagatagtcgagtcgcgtctctcgtcctcgtctgcctctgagaattccggaagaattc	869
Db	374	TTGAAACAGCTGGATTCGTCCGATCTGCTGGCCCGGAGACGTTCTCCACGAAAGG	433
QY	870	CGGATCTCTCGATGATGATGATCCTGAGCGAAAGTTCAGATATATACAGATATCCG	929
Db	434	ATGAGTCA-----TCGGAGATATTTTCAACAGTACCA-	465
QY	930	actcgatctactcgatataacgcgllcactctcttgagactctggacgggaattctctg	989
Db	466	-----TCGGATTACACTTCTTCAGATTTCCTCCACTCTGGACCGCAGTTCTCTG	514
QY	990	agaaactctctctgattctcactatctcaacattctgcgtctctgtactccagttcaag	1049

RESULT	11	612 bp	DNA	linear	GSS 13-DEC-2001
LOCUS	BH477964				
DEFINITION	BOHKJ17FR BOHK Brassica oleracea genomic clone BOHKJ17,				
ACCESSION	BH477964				
VERSION	BH477964.1				
KEYWORDS	GS:17686075				
SOURCE	Brassica oleracea.				
ORGANISM	Brassica oleracea				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.				
AUTHORS	Town,C.D., Van Aken,S., Uterback,T. and Fraser,C.M.				
TITLE	Whole genome shotgun sequencing of Brassica oleracea				
JOURNAL	Unpublished (2001)				
COMMENT	Other_GSSs: BOHKJ17FF Contact: Chris Town				
FEATURES					
source	7112 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: FR Class: sheared ends. Location/Qualifiers 1..612 /organism="Brassica oleracea" /strain="T01000DH3" /db_xref="taxon:3712" /clone="BOHKJ17" /clone_1lb="BOHK" /note="Vector: pHOST1. Site:1. BstXI; 2-3 kb sheared genomic DNA inserted into pHOST1 using BstXI linkers"				
BASE COUNT	198 a 113 c 94 g 207 t				
ORIGIN					
Query Match	4.38; Score 93.2; DB 12; Length 612;				
Best Local Similarity	67.48; Pred. No. 0.0015;				
Matches 225; Conservative 0; Mismatches 93; Indels 16; Gaps 6					
1769	gttagaacaacagatacagatgctgcgaatgcgttaagagctcgcagctgctgcgtgag 1828				
19	cttagggagagagattttaaatttgctcttcttattatcagagcgtcgatgcttcctggc 78				
1829	cagtaagcaatcaaaaagaacaaaccccta--aaacagagacacagctatct-cgata 1885				
79	cagtaagcaatcaaaaagaacaaacccccaacaggaacacagtaatctccacata 138				
1886	ccaacacacagagttatcattcattcaaaaacacacacagaglaaglaataagaaact 1945				
139	cgaaactcagaggttatattatccattta--ttaaacaatagatattatttgctaaact 197				
1946	-----cctctacagattatatactctaaacgcggcgcgagactaatlagctcttagat 1997				
198	aagcagatcccttcaacatttatattatatttcttgatc---cttgacattttttagctgttgatt 254				
1998	accacattatagtcgcccacattgctgcgtcatalacacattatcttcatcttccctaa 2057				
255	ggcactactaataatccacatttattgtgtcgcg-cacccaataattatgcttattttcccaa 313				
2056	tgcattagactctcatattctttaaagaatct 2091				
314	ttagcttagactctccattttttatttataatcatatttt 347				

[illegible]

TITLE	JOURNAL	COMMENT
Determination of clone and sequences from Entamoeba histolytica	HM1:IMSS sheared DNA library	Unpublished (2000)
Contact: Brendan J Loftus	Department of Eukaryotic Genomics	The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA	Tel: 301 838 0208	Fax: 301 838 3543
Email: b.loftus@tigr.org	Clones are derived from the Entamoeba histolytica HM1:IMSS sheared	DNA library
Seq primer: M13-Reverse	Class: Shotgun	High quality sequence start: 17
High quality sequence stop: 434.	Location/Qualifiers	1..791
1..791	/organism="Entamoeba histolytica"	/strain="HM1:IMSS"
/db_xref="taxon:5759"	/clone_lib="Entamoeba histolytica sheared DNA"	/note="Vector: pHD1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.	Genomic DNA isolated from broth cultures of E. histolytica	using a method described by Clark and Diamond (Clark,
C.G. and Diamond, I.S. (1997) Entamoeba histolytica: a	method for isolate identification. Exp. Parasitol.	77:450. The DNA was mechanically sheared to give a
light size distribution (~2 kb). The v + i method used for	the library construction is described in detail in Smith,	H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome	Sequencing: A Practical Approach, eds. M. Vaudin and B.	Bartell, Oxford University Press, 1999)."
BASE COUNT	275 a	78 c 45 g 393 t
ORIGIN		
Query Match	3.3%	Score 70.6; DB 12; Length 791;
Best Local Similarity	52.2%	Pred. No. 1.5;
Matches 157; Conservative	0; Mismatches 144; Indels 0; Caps 0;	
Oy 1844	agaagacacaaacccctaaacccgagcagcagatcctcgcgtacacacacgagttatca	1903
Db 790	AAAAATATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAA	731
Oy 1904	ttaactatcaacaaacacacacaggaagtaataagaaactccctacagattatata	1963
Db 730	TAAATTTTAAATATTAATAATATAATATAATATAATATAATATAATATAATATA	671
Oy 1964	cttaacgcagcgcgacttaattagctctagatcaccaattattagccacatttgg	2023
Db 670	TTTTTTTTTTTATTTTTTTTATTTATATATAATTTTTTTTTTTTTTTATATATAT	611
Oy 2024	tgcctacacacacattattcttattccctaaatcaatlagacctcatalatcitta	2083
Db 610	AAAAATATTTTAAATTTTTTTTTTTATTTTATTTTAAATTTATTTTAAATTTAA	551
Oy 2084	agaatattccctcgtctgaataaaaaaataaaaaaataaaaaaataaaaaaata	2143
Db 550	AAATATATATTTTTTTTTTAAATATAAAATATAAAATATAAAATATAAAATATA	491
Oy 2144	a 2144	
Db 490	A 490	
RESULT 14		
LOCUS	AL514901	668 bp mRNA linear EST 13-FEB-2001
DEFINITION	AL514901 LTL_NFL006.P12 Homo sapiens cDNA clone CLOB012ZH12 3	
	prime, mRNA sequence.	

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 05:59:50 : Search time 3869.65 Seconds
(without alignments)
11594.458 Million cell updates/sec

Title: US-09-821-839-1
Perfect score: 2144
Sequence: 1 actgcatcagccacactctct.....aaaaaaaaaaaaaaaaaaaaa 2144

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_bat:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_ju:*
20: em_mu:*
21: em_or:*
22: em_ov:*
23: em_pal:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description

C	1	1117.4	52.1	137699	8	AC006917	AC006917 Genomic s
C	2	78.4	3.7	5204	6	AX345803	AX345803 Sequence
C	3	78.4	3.7	34980	6	AX344555	AX344555 Sequence
C	4	70.6	3.3	4435	5	XM1MLZ	297214 Xenopus lae
C	5	68.6	3.2	149791	2	AC107226	AC107226 Oriza sat
C	6	67.4	3.1	6372	3	DDICMPA	M23449 Dictyosellu
C	7	67.2	3.1	34980	6	AX344558	AX344558 Sequence
C	8	65.6	3.1	7736	3	AF310889	AF310889 Dictyoste
C	9	65.6	3.1	34980	6	AX344553	AX344553 Sequence
C	10	65.6	3.1	34980	6	AX344554	AX344554 Sequence
C	11	65	3.0	34980	6	AX344574	AX344574 Sequence
C	12	63.4	3.0	545	6	AX185705	AX185705 Sequence
C	13	63.4	3.0	1090	3	AY069454	AY069454 Drosophil
C	14	63.4	3.0	1298	6	AX188346	AX188346 Sequence
C	15	63.4	3.0	6564	6	AX345025	AX345025 Sequence
C	16	63.2	2.9	1672	6	AX092286	AX092286 Sequence
C	17	62.8	2.9	268	6	AX186559	AX186559 Sequence
C	18	62.8	2.9	1730	9	AB063081	AB063081 Macaca fa
C	19	62.8	2.9	1838	10	BC012404	BC012404 Mus muscu
C	20	62	2.9	255	6	AX210988	AX210988 Sequence
C	21	61.8	2.9	34980	6	AX344564	AX344564 Sequence
C	22	61.6	2.9	1404	8	ATU62741	U62741 Arabidopsis
C	23	61.4	2.9	3715	10	MMU03846	U35846 Mus musculu
C	24	61.2	2.9	6535	6	AX345838	AX345838 Sequence
C	25	61.2	2.9	204652	2	PFMAL13P6	AL049183 Plasmodiu
C	26	61	2.8	113880	3	PFMAL3P4	AL008970 plasmodiu
C	27	60.8	2.8	425	6	AX284646	AX284646 Sequence
C	28	60.8	2.8	3392	3	AF300334	AF300334 Dictyoste
C	29	60.8	2.8	6944	6	AX347467	AX347467 Sequence
C	30	60.8	2.8	6944	6	AX349188	AX349188 Sequence
C	31	60.6	2.8	12177	6	AX345553	AX345553 Sequence
C	32	60.4	2.8	2497	9	AF124728	AF124728 Homo sapi
C	33	60.4	2.8	3999	3	LCU19108	Y19108 Lucilia cup
C	34	60.4	2.8	5179	6	AX323790	AX323790 Sequence
C	35	60.4	2.8	5179	6	AX346899	AX346899 Sequence
C	36	60.4	2.8	8945	6	AX345813	AX345813 Sequence
C	37	60.4	2.8	15714	6	AX346075	AX346075 Sequence
C	38	60.4	2.8	34980	6	AX344565	AX344565 Sequence
C	39	60.2	2.8	3363	3	CEMC81A	X75564 C.elegans m
C	40	60.2	2.8	8920	6	AX065855	AX065855 Sequence
C	41	60	2.8	1461	6	AX255406	AX255406 Sequence
C	42	60	2.8	1461	10	RNMAFA	X79812 R.norvegicu
C	43	60	2.8	6005	9	HSADENCR	Z35309 H.sapiens m
C	44	60	2.8	34980	6	AX344559	AX344559 Sequence
C	45	59.8	2.8	2744	3	DDSTAFTR	Y13097 D.discoideu

ALIGNMENTS

RESULT 3
AC006917/c
LOCUS AC006917
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome I, complete sequence.
ACCESSION AC006917
VERSION AC006917.6 GI:4757662
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana
thale cress.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 132699)
REFERENCE
AUTHORS Chao, O., Shinn, P., Dunn, P., Buehler, E., Kahn, S., Kim, C., Walker, M., Williams, S., Altafi, H., Araujo, R., Conn, L., Conway, A. B., Gonzalez, A., Hansen, N. F., Hutzar, L., Kremetska, I., Lenz, C., Li, J., Liu, S., Luros, S., Rowley, D., Schwartz, J., Toriumi, M., Vysotskaya, V., Yu, G., Davis, R. W., Federspiel, N. A., Theodis, A. and Ecker, J. R.
Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome I

TITLE
JOURNAL
Unpublished

REFERENCE 2 (bases 1 to 132699)
AUTHORS Ecker, J. R.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 132699)
REFERENCE Ecker, J. R.
AUTHORS Direct Submission
JOURNAL Submitted (07-MAY-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 132699)
REFERENCE Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q.,
AUTHORS Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bel, B., Chin, C.,
Chou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N.,
Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J.,
Liu, S., Mukharoky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwartz, J., Southwick, A., Thaver, A., Tortum, M., Vayenberg, M.,
Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th
Hamilton Walk, Philadelphia, PA 19104-6018, USA
On May 7, 1999 this sequence version replaced gi:4731042.
LOCATION/Organisms
1. 132699
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/complement(join(144..278,360..491,576..673,814..1435,
1749..1867,2161..2310,2441..2565,2698..2742,2837..2902,
2979..3031,3117..3234,3314..3451,3498..3655,3741..3818,
3894..3983,4054..4104,4138..4420))
/note="similar to Na/H antiporter dbj|BA09487.1"
/codon_start=1
/evidence=not_experimental
/product="F10B6.1"
/protein_id="AA079251.1"
/db_xref="GI:8778242"
/translation="MSTIGMALPYKSPERAIASSSSYSAENDSPVDAVTFAGTSLVL
GTACRYLNGTRVYVLLVIGFISGLGILIKLTVSSVSIWKLIEKGTKNL
GKLGIRINGINPDLIAVFLPILFESSFMDYQIRKGMQVLLAGGVLIST
FCIGALIKLFFPYNMKMTSLIGLIGATDPVAVVALLKEIGASKMTLLDGSIM
NDGAVVAVPOLFFKMGKNSDSTIKFLYONSGVAGLAFASVFKLFIIND
TVAOTVTLSSYRAYVYTAQDMANVSSILVMIKMPFAFATKAGDSLHFW
EMAKITANTLVFLSGVIAESVLSQITISYKVSOPILRPDFFLLFLDVGGN
GVAFLTLVYVQLSCVAVGVLPDLAKRSYGGLMKESIILWGLRAVLIQFLPLTG
GSAFLTVIRHSSIKDLKEDGROVPHDYGAGSIDPTNIMDIRLRLNISCVVSS
HCQTSYVKAIDSLRLPVSQCAWYEMLDRICTANVLMQSDALDIVSS
LSDMGLEPRVHPFNKYKFLQSKLIPKLVTHLIVERLESACYISSAFIRARIROO
LHIFGNSNINASTVINESEVEGEAKOFEDVSPQVSVLKTRQCHVYHLNNG
YTKNLEKGLLEGKESHLDVQSDLKLLRPPSLKLPNDVLLITSNPLKDNSSF
RSLAIGETDA"
/join(3650..5801,6081..6121,6230..6819)
/note="unknown protein; similar to EST emb|F14394.1"
/codon_start=1
/evidence=not_experimental
/product="F10B6.2"
/protein_id="AA079216.1"
/db_xref="GI:8778207"
/translation="MRTPTTLLVGAITFSGAGYRSDASHRKRECVLYANKV
GFHNPSETVYEDLPCTFEGVEKEKALGELANDRIVASAPYKLNFEDESEVYC
NKLISKREVKQFRKAVKDYFQWYDDLPIMGFIQKVDKDKSDPEKRYFLYKHIO
FEILYNDKRVLEISARDPHSLVDLEDEKVDADAFMYTAKMKETEPPEKREKYSMS
SSLPHLEIHFSTIISCVYLLTGFLATILRLVAKNDPMK"
/join(6881..7029,7140..7296,7392..7876,7961..8180)
/note="similar to transmembrane 9 superfamily member 2
gi|4758874: similar to dbj|AA042489.1, gb|AI96021.1,
emb|F14395.1"
/codon_start=1
/evidence=not_experimental
/product="F10B6.3"
/protein_id="AA079217.1"
/db_xref="GI:8778208"
/translation="MLKRYAODEADDOETGMVINGDVEFRTNSLFLASLSG
TOLFTLIFETFMALGVPEYKALFALVAVIALVSGIAGTSASFYCOLGKSGM
VRNLLTGCLGCGPLPFLFNTVAITATPAALPFGTIVYVILMTVLSPLVGS
GIKNSKAEFOAPCRFTKYPREIPLPYRNSAIPOMMAGFLPSAIVLELYFVS
VMGRITVYISLFIYFILLITATVATVLPOLAEEDHOMWRSTCGSGGLFI
YACIVYVYARSMSGFMOTSEFFGIMACTICGFFLMLCTVGFRAALLFVHITRSIK
GE"
8842..9714
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F10B6.4"
/protein_id="AA079218.1"
/db_xref="GI:8778209"
/translation="MPLSEVVDPIKETIDSLGPISDVLYQKLTVAESQRRLR
QTLVSRKMEKVIDLVSEASMSGLKFEVENKLSSEBDLVNOCIKWEKEC
FLVHODESLMEGNETDRARAEASRVELEEVKMSDEIKRISSECDLSIL
ASVSDESIISLGRIPLEANSBEDKCCFTLAKMDQLPTQKTVSLVSWKRIEKER
ECLIMNLAKAEQVEVLEVSQNEIDRENKFLROCSAERSHGMKFKRKSIMASSP
IEKRILESSQEPD"
11067..11906
/note="unknown protein; similar to EST gb|AI995472.1"
/codon_start=1
/evidence=not_experimental
/product="F10B6.5"
/protein_id="AA079219.1"
/db_xref="GI:8778210"
/translation="MDDDFRNMGGYEPAAATFKNLQLMSTIDRTNFKFLGRDP
NLMMGNGSYHNOEPIHNSYMWIINOOKRPNMLPVTTAPNGVNLPEPSSASMQ
MNHLHNOEBENKLEETIVOTKRRKNAKSGTPAKRDRPKDKNSSNNNT
NVRVLPKRSKNDLVKNGEIVNIOISGLPVDICGTAPOOCYRMCGGMOSSCTNS
MPLPSTKRKGRKIRISGRMSGARKKYLEKLASGPFNGNPIDLSKMHANHTNKEV
TIR"
complement(13762..14781)
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F10B6.6"
/protein_id="AA079250.1"
/db_xref="GI:8778241"
/translation="MKLKRRAVVLKDGPSLIADDDITAAVYATSHDELSIDTBSA
QFTRVYVSPSLKFLVLSISRKRTSRMAVLAIGMLHGFLLCKSTVAESITGL
PPLDSSFGBNRSKIMSGFNLFPVAFPLDRSLIFHDGNHRRNRESVLTIRLY
IIRKMOIYVDSLIRIKPIENNMIVPEAMENYVSEIMELYGMICRIEVLNPHVS
KIKRTADIALAKIVAKSMKGGLKRYEFCQDLGVSAQEIPIEVRIPEDVYIHDE
LVRTAASESESAERTEIEESEEDEETETKLSLITLDNEBAPASPPRVVVDIP
DLSP"
16027..16533
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F10B6.7"
/protein_id="AA079220.1"
/db_xref="GI:8778211"
/translation="MSTCYRECMRNNAKLSVAIDGCREYSPSTGDLVACGCH
RSYHRIDIVSSPQITNTRPPTSLRKYOLARKWKAKRENSREEDDPTLESTEEK
MYVORRRKSTFADEOREAKMDVAAKLGITLMDKALRREIIVREGICVYRIHRTYV
NNKRYH"
complement(join(17635..17815,17901..18019,18141..18194,
18300..18406,18547..18611,18742..18950))
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F10B6.8"
/protein_id="AA079249.1"
/db_xref="GI:8778240"
/translation="MEKAYFLLVWQEPRIQCEPVYVLFCTQWICLKKRQRIPLFW
YPEELHLODLPKQVIFVYFLIRIMFDFQVCKSSYTKGDSLVSGALGFNE
IVRYPVKSLGSEFKSLCPKSGSKVITGQALVGTSLTFVGEAVNRDKNALMIGSK
EGLSIVSESSFDEVMNMSQSBLCVILAKIGSIAVAIVYGVDAVKYLLPFW

QY	2036	attatctctcttctcccaattcatctagaccctcatctctaaagaatcttc	2095
Db	4304	ATTATTTTATTTTGGCTATGTATATATATTATCTTTTCACAAATAAAAGCTTCT	4365
QY	2096	tgtctgaaaaaataaaaaataaaaaataaaaaataaaaaa	2144
Db	4364	TTTTTAAAAAATAAAAAATAAAAAATAAAAAATAAAAA	4412

RESULT	5
AC107226/c	
LOCUS	
DEFINITION	
ACCESSION	AC107226
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
	149791 bp DNA lineat HTG 17-JAN-2002
	Oryza sativa chromosome 3 clone OSUNB0081P02, ** SEQUENCING IN
	PROGRESS ***, 10 ordered pieces.
	AC107226
	AC107226.1 GI:18182750
	HTG: HTGS_PHASE2.
	Oryza sativa.
	Oryza sativa.

	Matches	122: Conservative	0: Mismatches	89: Indels	0: Gaps	0:
OY	1581	ggtcttactctaaagctgtctgaagcaatccagaagttgtaanaagaaagcaaatcttgg				1640
Db	57654	GGTTTATCTGGAAGGCTGGGAATGACAGATGACAGCTGGCCAAATCTACTTGG				57595
OY	1641	ctgttaacctactatccgcgaacaaactcaactctgttttggccctcaactgttagcagctg				1700
Db	57594	CTTTGCTCTCAGCTTTCGGACACAAATACTCTCTCTTGGCCCTCCACTGTGGACGG				57535
OY	1701	cactcgtgtgttctgcgcctgcaclogaacacaacaaatctctgtcatcaacagagtcatata				1760
Db	57534	CAGTGGTATCCCTCTGCTTCCCTTGGCCACAAACAATGAGTCAATGCATTCGATTTGGTTATGG				57475
OY	1761	aggttcatgttagaacaacagataaagagct				1791
Db	57474	AGGTAATACTTTAATCTTCATCTACCTGGCT				57444

ORGANISM	RESULT	6
DDICNPA	DDICNPA	6372 bp DNA linear, INV 07-MAR-1995
LOCUS	Dichosellium discoidium cyclic nucleotide phosphodiesterase gene,	
DEFINITION	complete cds.	
ACCESSION	M23449.1	GI:695312
VERSION	M23449.1	GI:695312
KEYWORDS	cyclic nucleotide phosphodiesterase.	
SOURCE	D. discoidium (strain NC4; cell line AX3-K) DNA, clones pGP-1 and pC34.	
ORGANISM	Dichosellium discoidium	

Query Match	3.28;	Score 68.6;	DB 2;	Length 149791
Best Local Similarity	57.8%;	Pred. No. 1.1e-05;		

CDS
Join(4127..4146,4297..5635)
/EC_number="3.1.4.17"
/note="CDNA sequence is J02628"
/product="cyclic nucleotide phosphodiesterase"
/protein_id="AA63168.1"
/db_xref="GI:695313"
translation="MALNKLKILSLILPILINLVNHSOEDCDDEDIGISAENSE
RBSVNSDGSNENYNDYRPPMWNVYSGSFATKCRDASVYTIPIGTGGIGISENL
SEILTRKGSNLEFLALDAGTYMOCVRLLTRFKPNTLPNTTYSMAVLEPQRTSFLK
NHVSYFLGHSHLDHVGGLIVSPEDYLVANMDVOPPLNGGILIRKLEKPEIDPT
SSSILQKRTINGLSTINSISTLNQVPMNLPSEGRVYFSIASGIEVPELVPY
NATMSLVANEPSPSVKVPPELCHDNLISFLTDSISGQIAPSPDGVSSVAC
DWEGKIVAWKQIKIDKKAIFYETSPNPTPDAMGSHLRPDVAKLMDQLLVQSIQ
TSPMTNLKHVKLILHETIKPKQVADPMTQRYVYQQLKEANNNGVRITIPQSGPI
CI"

Inton 4147..4296
exon 4297..5606
/number=4
polya_signal 6015..6020
BASE COUNT 2614 a 749 c 504 g 2505 t
ORIGIN

Query Match 3.1%; Score 67.4; DB 3; Length 6372;
Best Local Similarity 51.1%; Pred. No. 1.8e-05;
Matches 158; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1833 aagcaatcaaaagaacaaacccctaaacacagagacagatatactccgataccaaac 1892
DB 5712 AAATAATAATAA 5771
QY 1893 acaggtatcatctattatcaaaacacacagagtaagtaagtaagtaagtaagtaag 1952
DB 5772 TAACTAAATATGAAACATATACCAAAAAAAAAAAAAAAAAAAAAAAAAAACATACGA 5831
QY 1953 agattatataacttaactgaagctggaacttaagctcttagtatccaattatagtc 2012
DB 5832 CAATTTTTTTTTTTTTTTTTTCTCTAAACAAATATACATATAATCATCTTAATAATAA 5891
QY 2013 caccatttgctgcctacatacattatcttcttccattccattacatgaactcga 2072
DB 5892 CTCCTCCTCCTTCCCTCTCTATATATATGCTTCAATGATGATGATGATGATGATGAT 5951
QY 2073 tatcttaaaagaataattctctgttgaaaaaataaataaataaataaataaataa 2132
DB 5952 TTTTAAAGTCAACCAACCTCTTTTAAACCAAAAAAAAAAAAAAAAAAAAAA 6011
QY 2133 aaaaaaaa 2141
DB 6012 AAAAAATAAA 6020

RESULT 7
AX344558/c AX344558 349980 bp DNA linear PAT 01-FEB-2002
LOCUS AX344558
DEFINITION Sequence 9 from Patent WO0200932.
ACCESSION AX344558
VERSION AX344558.1 GI:18492444
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (sites)
Olek,A., Piepenbrock,C. and Berlin,K.
TITLE
Diagnosis of known genetic parameters within the mhc
JOURNAL
Patent: WO 0200932-A 9 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
1..349980
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo

sapiens)-Original length of seq 1: 3.673778 <223>-split as
follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001
649.980-seq 03 600.001 949.980-seq 04 900.001
1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001
1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001
2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001
3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001
3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length
of seq 2: 3.673778 <223>-split as follows.-seq 14
0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16
600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001
1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001
2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001
2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001
3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001
3.673.778

BASE COUNT 97534 a 3966 c 76586 g 171894 t
ORIGIN

Query Match 3.1%; Score 67.2; DB 6; Length 349980;
Best Local Similarity 51.3%; Pred. No. 2.6e-05;
Matches 156; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1841 aaaaagaacaaaacccctaaacacagagacagatatactccgataccaaacaggtta 1900
DB 69065 AAAAAATATGTAACCGGAAACGAAACCTTAATAATAATCAACGACCACTACACT 69006
QY 1901 tcatctatcttaacaaacacacagagtaagtaagtaagtaagtaagtaagtaagtaag 1960
DB 69005 CCAACCTAATATACAAATATTAACCTCACCTCAAAAATAATATCAAAAATAATTAAC 68946
QY 1961 atactatcgagctgagcttaattagctcttgatataccaattatagtcgacacatt 2020
DB 68945 AACTATATATACGAAACCTCAATCCCAACTATTAATAAACTAAACAAAAAATCCG 68886
QY 2021 gctgcctacatacattatcttcttccattccattacatgaactcctcatatctta 2080
DB 68885 TTAACCTAATATTAACATTAATTAACGAAATCACTCACCTCACTCACTCACT 68826
QY 2081 aagaataattctcttgtaaaaaaataaataaataaataaataaataaataaataa 2140
DB 68825 ATTAACAAACAAATATTCATCTCAAAAATAAATAAATAAATAAATAAATAAATAA 68766
QY 2141 aaaa 2144
DB 68765 AAAA 68762

RESULT 8
AF310889/c AF310889 7736 bp DNA linear INV 06-MAR-2001
LOCUS AF310889
DEFINITION Dictyostelium discoideum RacD (racd) gene, complete cds, and
unknown genes.
ACCESSION AF310889
VERSION AF310889.1 GI:12007300
KEYWORDS
SOURCE
ORGANISM
Dictyostelium discoideum.
Dictyostelium discoideum.
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE
1 (bases 1 to 7736)
Rivero,F., Dislich,H., Glockner,G. and Noegel,A.A.
TITLE
The Dictyostelium discoideum family of Rho-related proteins
JOURNAL
Nucleic Acids Res. 29 (5), 1068-1079 (2001)
MEDLINE
21127961
REFERENCE
2 (bases 1 to 7736)
Rivero,F., Dislich,H. and Noegel,A.A.
TITLE
Direct Submission
JOURNAL
Submitted (02-OCT-2000) Institut fuer Biochemie I, Medizinische
Fakultaet, Universitaet zu Koeln, Joseph-Stelzmann-St. 52, Cologne
50931, Germany
FEATURES
Location/Qualifiers
1..7736

BASE COUNT	3205	a	1097	g	837	g	2597	t
mRNA								
CDS								
gene								
CDS								
polyA_signal								
mRNA								
CDS								

```

/organism="Dictyostellium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
join(<856..1007,1160..>4670)
/product="unknown"
/note="5' region identified manually, therefore unsure;
similarly to mouse SH3-domain containing adaptor protein
(central part) and rat L-glutamine aminohydrolase
(C-terminal part)"
/codon_start=1
/product="unknown"
/protein_id="AAG5120.1"
/db_xref="GI:12007301"
/translation="MFKMHQSNLVYVPFNIPANVTYKODTHAETTSSTIGLVKGFGL
NCEQIFKSKGAIKDDTVQYQMDQGVKWSHLSEITESSNAKAVVLEPATESE
EELSLPHOVDITISKDDSGMGEKSGKQNGVFPENFQVEPTVQPPEDDEASAEAT
KEELISLPIQVQSRIGINISIGSMDELKKLKYRTATYTSADI EAHNNMNAENES
NSSSSSSSSNNISGGEFVAVKPSGAVNQVQDQKVLNNROSTIIISGVPATISIP
SSNNTTTTTTTTTPAYQPLIQLKPKROTMEINTSGNSIPATSSQSTKESSSGGN
KSLTEFLPKLKTALTSALGSSNNESSGGDKNNSGKTRAKVYDFEKEENELN
LKQDIIIVLAKDSSGMQVGINSTQATWESNTEVELEKEKRLPPSSSENQPKL
LVKSPDSSSSSSQPOVVEAKLQSLDTPKELTARKAAGARKPPKSVRSAYLLSD
KARERETKQIEKPSHVYSSQKELSMYDNNSDELHLDQDDPSTPPISSRSGGFTFP
QOQOQOQVDRKIERPPSPPPPKRTAAMNSLNNINEPPSPNKPPPKSLKPLSTD
GYEPPSSSTSTPTTSGEDTLQKSPLOKRVAPQOQOEDLPSTLQKRPPLP
PTPKPRATQNNSESTAPAPPSHEVFNKMLPAKSSGPIGNSQELPTDQSSPNAT
TQPPPPPLKRVAKPPVQOQOQNSNENINIQPKLPSASPSTSSSPTTIDISTTK
PPAPRLKSKAAAPPPSPNEEATTTSPPTLAPKAPKALNRGSSAFTPSLOPOD
LPSTPHKRNLSNPTNNKPPPTKPKPOLPLPFOFNDICEVSNICNEIOESDUS
QETTTTKHNPSTLIDFFSVAVACSNDQGFQWQNDQLNDHCVPMPQVLEPPLYSILCE
BIDDEVSKFISDQAPRDSQDNQKPLNDQKSHNPFTLAGOLITANLSVNIQNHVLE
STQFNLIIQELVNCSSVCDMYLSQKETSDEILVGLKSSGLIKOPDDVLDPEY
QINISQLSKQSVLAAITANDGVCPPSQDKLAPRNITSKIDILRICNSTPTTSSI
FNHDEPKNNSKTIPIILSDSGVLMIIIPGLMGVSIISPQNNNNNNNNNNNFHYHFN
NNHIEPCKKSLTDDQ"
join(5116..5180,5293..5333,5523..5590,5805..5914,
6023..6634)
/gene="racd"
/note="Corresponds to EST clones SSB614, SSG550, SSJ204
and SSK792 from the Japanese cDNA sequencing project; cDNA
(parial) already described in J11594."
/product="racd"
/gene="racd"
/protein_id="AAG5121.1"
/db_xref="GI:12007302"
/translation="MATGIKRTYKVVVVGDDGAVKTSLLIITYTKAPKRDVYPTVFDN
FNCLEMDKPNVNLVMDTAGEBDYDNLRLSPQIDVFIICYSVAKDSLDNIKYK
LPEILNQTQGPPIIVGTQTDLRDKKTSIQLOESQOEFVSDEDAVALAKIGAVQFFE
ECSATGAGVNDIEPAALIKAFNKPATVPTSPSSGKSSPSSSTSSKPKKTTTTTSS
SSSSPASTAKPAGKEKKLSMGLFRKKDKKKPAK"
6612..6617
/gene="racd"
join(<7278..7533,7645..>7736)
/product="unknown"
join(7278..7533,7645..>7736)
/codon_start=1
/evidence="not_experimental"
/product="unknown"
/protein_id="AAG5122.1"
/db_xref="GI:12007303"
/translation="MSIOSRQDPFKPEELCKDPEYLPETITITDIIQYLPINPELIL
EKQSLSTKSTSYVAQVYIKGLKCVTSFNTSLTPQKDKMSNIMLITIEVQPLEQK
ALSLVFKLYSVLK"

```

Query Match	3.1%	Score 65.6;	DB 3;	Length 7736;
Best Local Similarity	51.0%;	Pred. No. 5e-05;		
Matches 155;	Conservative	0;	Mismatches 149;	Indels 0;
			Gaps	0;
QY 1841	aaaaagacacaaacccctaaacccgagacagatactccgatlaccacacacagtgta	1900		
Db 5348	AAAAAAATACATACCCACACACACATCACCTACACTACTACTTAAACAGTTTCTGT	5289		
QY 1901	tactactatttcaaaaaaagaacagagtaagtaataagaactccctccagatttt	1960		
Db 5288	TTTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	5229		
QY 1961	atactaatcgagctgagactaatatagctctatagataaccatattagtcgacacatt	2020		
Db 5228	TAAATCAACACATTTGTGTAAATTTATTTTGATATTAATCATTTACTTAATACCTGT	5169		
QY 2021	gtctcgctcaataacacttattcttatttcccaattcaatagactccatattccta	2080		
Db 5168	AGCCATATTTATTTTATTTTATTTTATTTTCTTTTATTTTATTTTAAAAAATAATGATGT	5109		
QY 2081	aaaaaatatttcttctgtttgaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	2140		
Db 5108	TAAAAAAATTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	5049		
QY 2141	aaaa 2144			
Db 5048	AAAA 5045			
RESULT 9				
AX344553/c				
LOCUS	AX344553	349980 bp	DNA	linear
DEFINITION	Sequence 4 from Patent WO0200932.			
ACCESSION	AX344553			
VERSION	AX344553.1	GI:18492439		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Olek, A., Piepenbrock, C. and Berlin, K.			
TITLE	Diagnosis of known genetic parameters within the mhc			
JOURNAL	Patent: WO 0200932-A 4 03-JAN-2002;			
FEATURES				
source				
	Location/Qualifiers			
	1. 349980			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="Chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.67378 <223>-split as follows. seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 follows. seq 03 600.001 949.980-seq 04 900.001 649.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.249.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3.673778 <223>-split as follows. seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"			
BASE COUNT	86882 a	5859 c	85073 g	172166 t
ORIGIN				
Query Match	3.1%;	Score 65.6;	DB 6;	Length 349980;


```

Db 251631 AAAAAATTATCTACGAGATATCTCAAAAAATTTCTTTCAACCAAAAAAACCTTCAATAC 251572
Qy 2070 tcatattcttaaaagaatattctctgttgaaaaaaagaaaaaagaaaaaagaaaaa 2129
Db 251571 AATATCACCTAAACCAAAAAATTAATCTATATCTCTAAAAAAAACCAAAAAATATTT 251512
Qy 2130 aaaaaaagaaaaa 2142
Db 251511 AATCAAAAAA 251499

RESULT 12
AX185705/c AX185705 545 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1400 from Patent WO0142467.
ACCESSION AX185705
VERSION AX185705.1 GI:15137111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 1400 14-JUN-2001;
Milemnum Predictive Medicine, Inc. (US)
FEATURES
source 1. 545
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 200 a 40 c 23 g 209 t 73 others
ORIGIN

Query Match 3.0%; Score 63.4; DB 6; Length 545;
Best Local Similarity 54.5%; Pred. No. 0.00015;
Matches 91; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 1978 actaatgccttagatgatacattatagtcaccattgtgcgcgcataacat 2037
Db 280 ANTTTNNCCCTNNNNNTTTNNNTTTTNNNTTTTNNNTTTTCCCCCAANTTTT 221
Qy 2038 tcatctcattatctccatcatcattagacccatcattcttaagaatattcctg 2097
Db 220 TTTTNNNNNTTTTNTTNTTNTTNNANTTTTNNAAAAAATTTTNTTNNANT 161
Qy 2098 ttgtaaaaaagaaaaaagaaaaaagaaaaaagaaaaa 2144
Db 160 TTTTNNAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 114

RESULT 13
AY069454 1090 bp mRNA linear INV 17-DEC-2001
LOCUS Drosophila melanogaster LD17744 full length cDNA.
ACCESSION AY069454
VERSION AY069454.1 GI:17862243
KEYWORDS FLI CDNA.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1090)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Abmayan,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Garfin,D., Frise,E., George,R.,
Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
Nuno,J., Paclob,J., Paragas,V., Park,S., Phonanavong,S., Wan,K.,
Yu,C., Lewis,S.E., Rubin,G.M. and Celisner,S.
TITLE Direct Submission

```

```

JOURNAL Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
location/Qualifiers
1. 1090
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="99D1-99D1"
1. 1090
/gene="CG7866"
/note="Alignment with genomic scaffold AE003771"
/db_xref="FlyBase:FBgn0039725"
39..500
/gene="CG7866"
/note="Longest ORF"
/codon_start=1
/db_xref="FlyBase:FBgn0039725"
/product="LD17744"
/protein_id="AAL39599.1"
/db_xref="GI:17862244"
/translation="MSIGNDMPHUKNTTKVKGICGSAANSRSPDNTPIPSNEEN
VQDKGKLEPPDSYLUKLEKRLARVQGNKLLNLRKRDCKRGLSSGVPISI
FQPLFEDATFESGRUHRHLPLVQAVNVEFIVEHDLQOASAEAOEE"
BASE COUNT 341 a 244 c 266 g 239 t
ORIGIN

Query Match 3.0%; Score 63.4; DB 3; Length 1090;
Best Local Similarity 66.4%; Pred. No. 0.00015;
Matches 91; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 2008 agtgcaccattgtgctcatcacattatcttattccctaattcattagac 2067
Db 939 AGAGTAACCTATTATGCTTANGTACATTCAGTATGCGACCCCTATTATTCATC 998
Qy 2068 tctcatctcttaagaatattcctgttgaaaaaaagaaaaaagaaaaa 2127
Db 999 TCAACACATCATTTAAACAAGTATTTAAGTCGCAAAAAAAGAAAAAAGAAAAA 1058
Qy 2128 aaaaaaagaaaaa 2144
Db 1059 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1075

RESULT 14
AX188346 1298 bp DNA linear PAT 06-AUG-2001
LOCUS Drosophila melanogaster LD17744 full length cDNA.
ACCESSION AX188346
VERSION AX188346.1 GI:15139819
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE
1 (bases 1 to 1298)
Barron A and Zhao Y
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

[illegible]

2030 atacacattattcttatttccctaattcattagacccccaattcctaadaagaa

1 FILING DATE: 06-APR-1995
2
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: IL 109257
5
6 FILING DATE: 08-APR-1994
7
8 ATTORNEY/AGENT INFORMATION:
9 NAME: BROWDY, Roger L.
10
11 REGISTRATION NUMBER: 25,618

RESULT 6
 US-08-462-831-9
 : Sequence 9, Application US/08462831
 : Patent No. 5552142
 :
 GENERAL INFORMATION:
 :
 APPLICANT:
 :
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
 :
 TITLE OF INVENTION: DERMATOPHAGOIDES
 :
 NUMBER OF SEQUENCES: 13
 :
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 :
 STREET: 60 STATE STREET, SUITE 510
 :
 CITY: BOSTON
 :
 STATE: MA
 :
 COUNTRY: USA
 :
 ZIP: 02109
 :
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 :
 COMPUTER: IBM PC compatible
 :
 OPERATING SYSTEM: PC-DOS/MS-DOS
 :
 SOFTWARE: ASCII TEXT
 :
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,831
 :
 FILING DATE:
 :
 CLASSIFICATION: 424
 :
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/945,288
 :
 FILING DATE: 10 SEPTEMBER 1992
 :
 APPLICATION NUMBER: US 580,655
 :
 FILING DATE: 11 SEPTEMBER 1990
 :
 APPLICATION NUMBER: US 458,642
 :
 FILING DATE: 13 FEBRUARY 1990
 :
 ATTORNEY/AGENT INFORMATION:
 NAME: MANDRAGOURAS, AMY E.
 :
 REGISTRATION NUMBER: 36,287
 :
 REFERENCE/DOCKET NUMBER: IPC-0100C (INT-024)
 :
 TELECOMMUNICATION INFORMATION:
 :
 TELEPHONE: (617) 227-7400
 :
 TELEFAX: (617) 227-5941
 :
 INFORMATION FOR SEQ ID NO: 9:
 :
 SEQUENCE CHARACTERISTICS:
 :
 LENGTH: 1172 base pairs
 :
 TYPE: nucleic acid
 :
 STRANDEDNESS: single
 :
 TOPOLOGY: linear
 :
 MOLECULE TYPE: CDNA
 :
 FEATURE:
 :

RESULT 10
 US-09-092-770-8
 : Sequence 8, Application US/09092770
 : Patent No. 5973119
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Coats, Steven R.
 : APPLICANT: Bass, Michael B.
 : APPLICANT: Robinson, Murray O.
 : TITLE OF INVENTION: NO. 5973119el Cyclin E Genes and Proteins
 : FILE REFERENCE: A-524
 : CURRENT APPLICATION NUMBER: US/09/092.770

Query Match	2.78;	Score 57.4;	DB 2;	Length 1215;
Best Local Similarity	30.68;	Pred. No. 2c-05;		
Matches 112;	Conservative 68;	Mismatches 180;	Indels 6;	Gaps 1;

```

RESULT 11
US-09-222-851-8
: Sequence 8, Application US/09222851
Default US, C107753

```

Query Match	2.7%	Score 57.4	DB 4	Length 1215
Best Local Similarity	30.6%	Pred. No. 2e-05		
Matches 112	Conservative 68	Mismatches 180	Indels 6	Gaps 1

QY	1233	atactgtgtccatccctcgtctacgtccatccacggttcacatgatgttaagaacatggt	1292
b	401	aysnaayttnarconcaratamnnwaaahvnyvnyvnyvnyvnyvnyvnyvnyvnyv	450

```
QY 1293 ctgacatggggttcacgaagaagacatgttcttaggagttgttctgttgatcgtctcc 1352
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 argntayacnynlcaayngnagaracnttytaytngcncaargtytlytgaymgtlyta 520
QY 1353 tgagcaagagatcattcaaacgcaagagctcctaactacatcagatcgagatcgagctta 1412
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 521 tgytaacncaatagayathaaayaaatagatgyncaatgncatgnaatncaawstnt 580
QY 1413 ctctggccaccagaatctgaagaaatcaacattcaacagatccgaaagaaagaaattca 1472
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 tathcnwnaarytngargarathatgcncaaa-----rytncargatgycont 634
QY 1473 ccattcagacctaagaatccgcatcggaagtgtgagcaatgagtgctgctgctcaag 1532
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 aytacnagayagngcnlygwsngargargayathymgnatgagatynatnathtyna 694
QY 1533 aagtcctcaactcaatgcttcacaccacaatctcaactctctgtgtgttactta 1592
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 695 argcnytnaargtgargatngtgcngtnacnathatwntgynaaytnttytnc 754
QY 1593 aagctg 1598
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 755 argtng 760

RESULT 12
US-09-342-681C-12
; Sequence 12, Application US/09342681C
; Patent No. 6355782
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 52978
; CURRENT APPLICATION NUMBER: US/09/342,681C
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3720
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (260)..(1506)
; NAME/KEY: misc.feature
; LOCATION: (2961)..(3673)
; OTHER INFORMATION: n represents a, c, t, or g
US-09-342-681C-12

Query Match 2.7%; Score 57; DB 4; Length 3720;
Best Local Similarity 70.8%; Pred. No. 4.3e-05;
Matches 75; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 2037 ttattcttattcttcctaatcattagactcgaatctcctcaattcttaagaatattcctt 2096
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3615 tatataatttcttctatctatattattattattattattattattattattattattatt 3674
QY 2097 gtttgaataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaa 2142
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3675 aaagggaataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaa 3720

RESULT 13
US-09-111-730-5
; Sequence 5, Application US/09111730
; Patent No. 6274359
; GENERAL INFORMATION:

Query Match 2.6%; Score 56.6; DB 4; Length 2469;
Best Local Similarity 60.9%; Pred. No. 4.5e-05;
Matches 92; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
APPLICANT: Hideharu Anazawa
APPLICANT: Hiroko Shimada
APPLICANT: Seiji Sugimoto
APPLICANT: Toshimasa Shinku
APPLICANT: Tatsuo Suda
APPLICANT: Yuzuru Ishimura
APPLICANT: Matsuhiko Hayashi
APPLICANT: Toshiaki Morikawa
APPLICANT: Tadaaki Yoshida
APPLICANT: Shu Makino
APPLICANT: Takao Saruta
APPLICANT: Hiromichi Suzuki
; TITLE OF INVENTION: 25-HYDROXYVITAMIN D3-1a-HYDROXYLASE AND DNA ENCODING THE HYDRO
; FILE REFERENCE: 1074
; CURRENT APPLICATION NUMBER: US/09/111,730
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24)..(1526)
US-09-111-730-5
```

```
Query Match 2.6%; Score 56.6; DB 4; Length 2469;
Best Local Similarity 60.9%; Pred. No. 4.5e-05;
Matches 92; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 1993 agttaccattttggtgcaccatttgytgcgtcatagacattattcttatttctc 2052
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2319 agtggaccacttactctccaccattcttcgacccacccgtcttattat 2378
QY 2053 cctaattcattagactcattcattcttaaaagaatattcttctgtttgaaaaaa 2112
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2379 ttaaaatgtatttatttattgtttccaataaatgtttactcttgaaaaaa 2438
QY 2113 aaaaaaaatataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaa 2143
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2439 aaaaaaaatataaaaaaaatataaaaaaaatataaaaaaaatataaaaaaa 2469
```

```
RESULT 14
US-09-248-335-51
; Sequence 51, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,759
; PRIOR FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 51
; LENGTH: 967
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-51
```

```
Query Match 2.6%; Score 56.4; DB 3; Length 967;
Best Local Similarity 68.4%; Pred. No. 3.3e-05;
Matches 78; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2031 taccattatttcttattcttccttaattcattagatctcattcttaaaagaatat 2090
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 06:59:30 ; Search time 339.91 seconds
(without alignments)
10829.527 Million cell updates/sec

Title: US-09-821-839-1
Perfect score: 2144
Sequence: 1 actgcatcagccactctct.....aaaaaaaaaaaaaaaa 2144

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq-037802.*

- 1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
- 23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2144	100.0	2144	22	AAD21310
2	1117.4	52.1	3970	22	AAD21312
3	123	5.7	3018	22	AAD21311
4	78.4	3.7	5204	24	ABL32901
5	63.4	3.0	585	22	AAH70126
6	63.4	3.0	1288	22	AAH72764
7	63.4	3.0	6564	24	ABL32123
8	63.2	2.9	1672	22	AAF92066
9	63.2	2.9	1701	21	AA64982

10	63.2	2.9	1701	22	AAF44128	Human PRO4063 (UNO
11	62.8	2.9	268	22	AAH70080	Human cervical can
12	62	2.9	216	22	AAK8314	Human digestive sy
13	62	2.9	216	22	AAI5755	Human colorectal c
14	62	2.9	255	22	AAH82121	Rat differential t
15	62	2.9	705	22	AAH33177	Human colon cancer
16	61.8	2.9	246	22	AAH33177	Human colon cancer
17	61.2	2.9	2115	22	AAH29131	CDNA encoding for
18	61.2	2.9	6535	24	ABL32936	Human breast cance
19	60.8	2.8	292	19	AAV21153	Human immune syste
20	60.8	2.8	425	22	AAH60450	Human immune syste
21	60.6	2.8	12177	24	ABL32651	Human cancer agent
22	60.4	2.8	337	22	AAH11578	Human immune syste
23	60.4	2.8	1267	21	AACT7794	Human breast cance
24	60.4	2.8	2497	21	AAH61197	Human cancer assoc
25	60.4	2.8	5179	24	ABL33997	Human CDNA encodin
26	60.4	2.8	8945	24	ABL32911	Human immune syste
27	60.4	2.8	15714	24	ABL33173	Human immune syste
28	60.2	2.8	8920	15	AAH62924	Human immune syste
29	60	2.8	1461	16	AAH01471	Cardiac myl-phosphat
30	60	2.8	1461	16	AAH01471	Mammalian mast cel
31	59.8	2.8	7722	12	AAH012023	Rat mast cell func
32	59.8	2.8	7753	15	AAH056826	Normal virus geno
33	59.6	2.8	443	8	AAH70773	Sequence encoding
34	59.6	2.8	3302	22	AAH34933	Human colon cancer
35	59.6	2.8	3380	22	AAH81743	Human differential
36	59.4	2.8	763	22	ABL08093	Human breast cance
37	59.4	2.8	5153	24	ABL2955	Human immune syste
38	59.4	2.8	5464	24	ABL33916	Human immune syste
39	59.4	2.8	11835	22	AAH45394	Chemically pretrea
40	59.2	2.8	6070	24	ABL33679	Human immune syste
41	59.2	2.8	6070	24	ABL33679	Human metastasis a
42	59.2	2.8	567	22	AAH89704	Human adult T-cell
43	59	2.8	7345	24	ABL32345	Human immune syste
44	58.8	2.7	612	22	AAH71471	Human cervical can
45	58.6	2.7	243	22	AAH29057	CDNA encoding for

ALIGNMENTS

RESULT 1	
ID AAD21310	standard: CDNA: 2144 BP.
XX	
AC AAD21310:	
XX	
DT 28-JAN-2002 (first entry)	
XX	
DE Arabidopsis thaliana SDS cDNA.	
XX	
KW Meiotic prophase I; chromosome 1; male sterile SDS mutant; apomixis;	
KW Plant breeding; seed production; SDS protein; ss.	
XX	
OS Arabidopsis thaliana.	
XX	
FT Key	Location/Qualifiers
FT CDS	98..1834
FT	/tag="a
FT	/product="SDS protein"
XX	
PN WO200174144-A1.	
XX	
PD 11-OCT-2001.	
XX	
PF 29-MAR-2001; 2001WO-US09875.	
XX	
PR 31-MAR-2000; 2000US-193523P.	
XX	
PA (TYPE-) UNTV PENNSYLVANIA STATE.	
XX	
PI Ma H;	
XX	

DR WPI: 2001-662939/76.
XX P-SDB: AAE12998.
XX A new plant gene from Arabidopsis, designated SDS, mutations in which
PT are associated with inability to produce pollen, is useful for the
PT production of male sterile plants for plant breeding -
XX
XX Claim 8; Page 39; 47pp; English.
XX The patent discloses novel plant gene from Arabidopsis thaliana,
CC designated SDS, which is associated with a failure to maintain,
CC homologous attachment during meiotic prophase I. The SDS gene is
CC located on chromosome 1 and is useful in plant breeding to produce
CC male sterile SDS mutants and cloned progeny by apomixis where
CC meloisis is bypassed in seed production. The gene may also be
CC used as a probe to identify related genes in other plant species,
CC and to identify and isolate other genes of the meloisis regulatory
CC pathway. The present sequence is a cDNA encoding Arabidopsis thaliana
CC SDS protein.
XX
SQ Sequence 2144 BP; 649 A; 461 C; 467 G; 567 T; 0 other;
Query Match 100.0%; Score 2144; DB 22; Length 2144;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 acgagcatagccacactctagctctctgaactaagcaactcttccaatcgaatt 60
DB 1 acgagcatagccacactctctgaactaagcaactcttccaatcgaatt 60
QY 61 tctaatcttagttcaagcttcgtaagaaaaaagaagagatcgcgtatgagaa 120
DB 61 tctaatcttagttcaagcttcgtaagaaaaaagaagagatcgcgtatgagaa 120
QY 121 ttaagaagcaagcctcgagacgacgcgttcgscggaagaagctccggtcgaagcatt 180
DB 121 ttaagaagcaagcctcgagacgacgcgttcgscggaagaagctccggtcgaagcatt 180
QY 121 ttaagaagcaagcctcgagacgacgcgttcgscggaagaagctccggtcgaagcatt 180
DB 121 ttaagaagcaagcctcgagacgacgcgttcgscggaagaagctccggtcgaagcatt 180
QY 181 acgcggaagagagacagacatctccggttccttgtaacacactctctggaacaaca 240
DB 181 acgcggaagagagacagacatctccggttccttgtaacacactctctggaacaaca 240
QY 241 aatcgaagctctgctgctctgctgcatcctgctccgacttggttaagctatgacaact 300
DB 241 aatcgaagctctgctgctctgctgcatcctgctccgacttggttaagctatgacaact 300
QY 301 tctcgtggtctcgagacagatcgaagagctcgatccgaagaagactctaaattgaga 360
DB 301 tctcgtggtctcgagacagatcgaagagctcgatccgaagaagactctaaattgaga 360
QY 361 ggttagaagcttctaaccgtgtatacatgttaagagagacgtatgtatccgaatttcg 420
DB 361 ggttagaagcttctaaccgtgtatacatgttaagagagacgtatgtatccgaatttcg 420
QY 421 aaggaatcagaggtcttactctaaagctacaagaagagagagagatgagatcgaagt 480
DB 421 aaggaatcagaggtcttactctaaagctacaagaagagagagagatgagatcgaagt 480
QY 481 aaggaatcagaggtcttactctaaagctacaagaagagagagagatgagatcgaagt 540
DB 481 aaggaatcagaggtcttactctaaagctacaagaagagagagagatgagatcgaagt 540
QY 541 gggaaataaaatlaaacgacaagatgagatctcttctcagatcgaatgtgacactcgc 600
DB 541 gggaaataaaatlaaacgacaagatgagatctcttctcagatcgaatgtgacactcgc 600
QY 601 cggagacatgctcacaacacgaggtttgaattcgaatcggagataagagagcgacgt 660
DB 601 cggagacatgctcacaacacgaggtttgaattcgaatcggagataagagagcgacgt 660
QY 661 cgttctctgcatatctgagatgagatcttccaagctcggagacgttaccggaagagc 720
DB 661 cgttctctgcatatctgagatgagatcttccaagctcggagacgttaccggaagagc 720

DB 661 cgttctctgcatatctgagatgagatcttccaagctcggagacgttaccggaagagc 720
QY 721 tgatacgaagaatattgaatctcccaagcgaagacgtctgtgaaagctgattccctctc 780
DB 721 tgatacgaagaatattgaatctcccaagcgaagacgtctgtgaaagctgattccctctc 780
QY 781 tgaatcgcgcgaagaatgaaagcgcgagcttgagatagtcggaatgacgtcgtatcgc 840
DB 781 tgaatcgcgcgaagaatgaaagcgcgagcttgagatagtcggaatgacgtcgtatcgc 840
QY 841 tgcctcgaagaatctcgaagaaggttcggaattctctcgaatgagatcgcgtacga 900
DB 841 tgcctcgaagaatctcgaagaaggttcggaattctctcgaatgagatcgcgtacga 900
QY 901 acgtcagaagatataatcacagatctccgaactcgaattcgaatcgaatcgcgtacat 960
DB 901 acgtcagaagatataatcacagatctccgaactcgaattcgaatcgaatcgcgtacat 960
QY 961 ctctctgacacttgcgacgagatctctggaatctctcctgattctcctatttaca 1020
DB 961 ctctctgacacttgcgacgagatctctggaatctctcctgattctcctatttaca 1020
QY 1021 tctcgcctctgtacacccagctcgaagacagctctgtagatccacgattcccaaga 1080
DB 1021 tctcgcctctgtacacccagctcgaagacagctctgtagatccacgattcccaaga 1080
QY 1081 ttttgatcttctgagagagagaaattcaactctgaaattgtaaggcttgatgagagga 1140
DB 1081 ttttgatcttctgagagagagaaattcaactctgaaattgtaaggcttgatgagagga 1140
QY 1141 ggtggaagagacatctcaagagcgaagagaaagaagaagacatgacatataatcgga 1200
DB 1141 ggtggaagagacatctcaagagcgaagagaaagaagaagacatgacatataatcgga 1200
QY 1201 ctggtcgaagacatctcgcagagatgagacatctggtctcaatctcctcgtctacgctc 1260
DB 1201 ctggtcgaagacatctcgcagagatgagacatctggtctcaatctcctcgtctacgctc 1260
QY 1261 catcagttcaatgagatctgaaagacatctctgacagagggcctcagcaagaagacatt 1320
DB 1261 catcagttcaatgagatctgaaagacatctctgacagagggcctcagcaagaagacatt 1320
QY 1321 gttctcagagatggtctgtgtgataagatctcccgaaacaaagatcatctcaaaagcgaag 1380
DB 1321 gttctcagagatggtctgtgtgataagatctcccgaaacaaagatcatctcaaaagcgaag 1380
QY 1381 gactcctaactagtcggaatctcgagcttaactctgacacacagaattgagaatataca 1440
DB 1381 gactcctaactagtcggaatctcgagcttaactctgacacacagaattgagaatataca 1440
QY 1441 acccttaacacagacatccggaagaagaaattcaacattgaaagctatagacagga 1500
DB 1441 acccttaacacagacatccggaagaagaaattcaacattgaaagctatagacagga 1500
QY 1501 tgaagtggtgagcaatgagatggtctgttcaagaagctcctcaacttcaaatgtctacacc 1560
DB 1501 tgaagtggtgagcaatgagatggtctgttcaagaagctcctcaacttcaaatgtctacacc 1560
QY 1561 cacaactctcaactctgtgtgtcttaacttaaaagctgttcgaagcaatccgaagttga 1620
DB 1561 cacaactctcaactctgtgtgtcttaacttaaaagctgttcgaagcaatccgaagttga 1620
QY 1621 aaggaagcacaatccttgcgttcaactcaatccgacaaactcaactcgtttttg 1680
DB 1621 aaggaagcacaatccttgcgttcaactcaatccgacaaactcaactcgtttttg 1680
QY 1681 ggcctcaactgttagcagctgacgtctggttctcgcgtgcatcgacaacaaatctc 1740
DB 1681 ggcctcaactgttagcagctgacgtctggttctcgcgtgcatcgacaacaaatctc 1740
QY 1741 tgcatacacaagatctcaaaagttcaatgttgaagaacaaagataacaggttgcctgagatg 1800
DB 1741 tgcatacacaagatctcaaaagttcaatgttgaagaacaaagataacaggttgcctgagatg 1800

Query Match	Similarity	52.1%	Score 1117.48	DB 22	Length 3970
Best Local	Similarity	99.9%	Pred. No. 4,1e-218		
Matches 1118	Conservative	0	Mismatches 1	Indels 0	Gaps 0
QY	1	actgcatacagccacatctctctagctctcgtacctaacgaactccatatttcaaaatcgaatt	60		
Db	510	actgcatacagccacacatctctagctctctctacaaagaactccatatttcaaaatcgaatt	569		
QY	61	tctaattctcagtttaacgaactctctgtacggagaaaataatgaagagatctcgatctggaga	120		
Db	570	tctaattctcagtttaacgaactctctgtacggagaaaataatgaagagatctcgatctggaga	629		
QY	121	ttcaaacgcgaagcctctgagacgcagcgttctgcgcggagaagaagctccggtctgacggaatt	180		
Db	630	ttcaaacgcgaagacccctgagacgcagcgttctgcgcggagaagaagctccggtctgacggaatt	689		
QY	181	acgcgcgaagaagagacagatctctcccggttctcttcaatacactctctcggagcaaaa	240		
Db	690	acgcgcgaagaagagacagatctctcccggttctcttcaatacactctctcggagcaaaa	749		
QY	241	aatcgcgaagctctcgcgtcttctgtcgaattcttcgtctccgaattgttgatctataaacaaact	300		
Db	750	aatcgcgaagctctcgcgtcttctgtcgaattcttcgtctccgaattgttgatctataaacaaact	809		

Db 810 ttccctggttcgacgaatcgagaagcgcgaatccgaagaagactctaatcgaga 869

— — — — —

[illegible][illegible][illegible][illegible]

541 gggaaataaaattaaagacaaacgaatgaatctctttctcagaatccgaatgaaccttcac 600

Db 1050 qqqaataaatttacqacacacqatqaadatctcttctcacqatccqatqtgaacctgc 1109

QY 601 cggacatgtctccaacagccggaqtttgaatttcgaatcggagaataaaggagcgacgt 660

Db 1110 cgacatgtctccaacagccggagtttgaaatttcgaatcggagaataagagagcgacgt 1169

661 cgttctgtcatatctgga gttgagtactgttccaagltcgggagcgttaccgga ggaac 720

Db 1170 cgtttctgtacatctctggaggttgagtaactgtttccaagttcgggagcgcttacccggagggagc 1229

QY 721 tgataacgaagaattgaatctccaagccgagcagcttcgtggaagctgattcctct 780

Db 1230 tgataacgaagaattgaatatctccaagccgacgactctgtagaagctgattcctctc 1289

QY 781 tggatcggccaaggaattgaa gccgagcttgagatagtcggaatgcctctctgatctgc 840

Db 1290 tggatcggccaagaaattgaagccggaagcttgagatagtcggaatgcgtctctgatactcgc 1349

841 ttgctctgagaaatctctcggaagaggtcttcggaatctcttcgacatcgagagccatctcgagca 900

DD 1330 LRGRLRGYAAAA LRLRLRGYAAAYAGY LRLRLRYALRLRLRLRYALRYAGY LCAALRYAGY CA

201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 117

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Db 2327 CCAATATCCACCAAAAATACACATTAATATTTTTCACGTAACGCCCTA 2268
OY 2073 tatcttaaaagaatattctctgttgaaaaaaaaaaaaaaaaaaaaaaa 2132
Db 2267 ATAAACAAAAAATATATATATATATATATATATATATATATATATAT 2208
OY 2133 aaaaaaaaaa 2144
Db 2207 AACCAAAATTA 2196

RESULT 5
AAH70126/c
ID AAH70126 standard; cDNA; 545 BP.
XX
AC AAH70126;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 1400.
XX
KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Deeds J, Berger A, Zhao X;
PT WPI: 2001-375006/39.
XX
DR New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
PS Claim 1; Page 319-320; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 545 BP; 200 A; 40 C; 23 G; 209 T; 73 other;

Query Match 3.0%; Score 63.4; DB 22; Length 545;
Best Local Similarity 54.5%; Pred. No. 0.00092;
Matches 91; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 1978 actaatagctcttagtaacaaattatagtcaccattgtgtgcctacatcat 2037
Db 280 AANNTTNNCCCTNNNANNTTNNANNTTNTTNTTNNANNTTNTTCCCCCAANTTTT 221
OY 2038 ttattctatttccctaatcattagactcattctcttaaaagaatattcctgt 2097
Db 220 TTTTNNNNNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 161

OY 2098 ttgaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2144
Db 160 TTTTNNAAAAAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAA 114

RESULT 6
AAH72764
ID AAH72764 standard; cDNA; 1298 BP.
XX
AC AAH72764;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4038.
XX
KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PI Schlegel R, Deeds J, Berger A, Zhao X;
PT WPI: 2001-375006/39.
XX
DR New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
PS Claim 1; Page 838-839; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 1298 BP; 507 A; 201 C; 373 G; 215 T; 2 other;

Query Match 3.0%; Score 63.4; DB 22; Length 1298;
Best Local Similarity 75.2%; Pred. No. 0.0011;
Matches 79; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 2040 attcttatttccctaatcattagactcattcattcttaaaagaatattcctgtt 2099
Db 735 attactgacttctcttgagtagtttctctgaaatcattaaagtgatttcttactt 794
OY 2100 tgaaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2144
Db 795 taanaaaaaaAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 839

RESULT 7
ABL32123/c
ID ABL32123 standard; DNA; 6564 BP.
XX


```
Db 1543 ttctctgtctaattggaagataactcatctttaataaattatgtctaagataaaa 1602
Oy 2105 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2144
    |||||
Db 1603 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1642

RESULT 9
AA264982
ID AA264982 standard; cDNA; 1701 BP.
XX
AC AA264982;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1063 encoding cDNA.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunodhesin; gene mapping; ss.
XX
OS Homo sapiens.
XX
PM WC9963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99MO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088217.
PR 05-JUN-1998; 98US-0088218.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088732.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089030.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 17-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089539.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.

PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091526.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 09-JUL-1998; 98US-0091982.
PR 10-JUL-1998; 98US-0092172.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094551.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
```

PR	18-AUG-1998;	98US-0096960.
PR	18-AUG-1998;	98US-0097022.
PR	19-AUG-1998;	98US-0097141.
PR	20-AUG-1998;	98US-0097218.
PR	24-AUG-1998;	98US-0097661.
PR	26-AUG-1998;	98US-0097951.
PR	26-AUG-1998;	98US-0097952.
PR	26-AUG-1998;	98US-0097954.
PR	26-AUG-1998;	98US-0097955.
PR	26-AUG-1998;	98US-0097971.
PR	26-AUG-1998;	98US-0097974.
PR	26-AUG-1998;	98US-0097978.
PR	26-AUG-1998;	98US-0097979.
PR	26-AUG-1998;	98US-0097986.
PR	31-AUG-1998;	98US-0098014.
PR	16-SEP-1998;	98US-0098525.
PR	12-JAN-1999;	98US-0100634.
XX		99US-0115565.
PA	(GETH) GENENTECH INC.	
XX		
P1	Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;	
P1	Wood WI, Yuan J;	
DR	WPI: 2000-072883/06.	
DR	P-PSDB: AAY66654.	
XX		
PT	Membrane-bound proteins and related nucleotide sequences -	
XX		
PS	Claim 2; Fig 65; 822pp; English.	
CC	The invention provides membrane-bound PRO polypeptides and	
CC	polynucleotides encoding them. The PRO sequences of the invention were	
CC	identified based on extracellular domain homology screening. The PRO	
CC	sequences have homology with proteins including LDL receptors, TIE	
CC	ligands and various enzymes. The membrane-bound proteins and receptor	
CC	molecules are useful as pharmaceutical and diagnostic agents. Receptor	
CC	immunoadhesins, for instance, can be used as therapeutic agents to block	
CC	receptor-ligand interactions. The membrane-bound proteins can also be	
CC	employed for screening of potential peptide or small molecule inhibitors	
CC	of the relevant receptor/ligand interaction. The PRO encoding sequences	
CC	are useful as hybridization probes, in chromosome and gene mapping and in	
CC	the generation of antisense RNA and DNA. PRO nucleic acid sequences	
CC	will also be useful for the preparation of PRO polypeptides, especially	
CC	by recombinant techniques.	
XX		
SQ	Sequence 1701 BP; 557 A; 288 C; 381 G; 475 T; 0 other;	
Query Match	2.9%; Score 63.2; DB 21; Length 1701;	
Best Local Similarity	77.0%; Pred. No. 0.0012;	
Matches 77; Conservative	0; Mismatches 23; Indels 0; Gaps	0
OY	2045 ttatttcctcaatcattgagactgcataattcctaagaagaatatcttggttgaaa	2104
Dy	1572 tttctcttgtaattggagaagttaacctattttaaaaaatcatgctaagtacaaa	1631
OY	2105 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	2144
Dy	1632 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	1671
RESULT 10		
ID	AAF44128 standard; cDNA: 1701 BP.	
XX		
AC	AAF44128;	
XX		
DT	02-APR-2001 (first entry)	
XX		
DE	Human PR04063 (UNQ128) nucleotide sequence SEQ ID NO:114.	
XX		
NM	Human: secreted and transmembrane protein; PRO: cytosolic;	

cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
diagnostic assay; ss.

Hom sapiens.

MO200073454-A1.

07-DEC-2000.

30-MAR-2000; 2000MO-US08439.

02-JUN-1999; 99MO-US12252.
23-JUN-1999; 99US-014103.
07-JUL-1999; 99US-0143048.
20-JUL-1999; 99US-0144758.
26-JUL-1999; 99US-0145698.
28-JUL-1999; 99US-0146222.
17-AUG-1999; 99US-0149396.
15-SEP-1999; 99MO-US21090.
15-SEP-1999; 99MO-US21547.
08-OCT-1999; 99US-0158663.
30-NOV-1999; 99MO-US28313.
01-DEC-1999; 99MO-US28301.
16-DEC-1999; 99MO-US30095.
20-DEC-1999; 99MO-US30911.
05-JAN-2000; 2000MO-US00219.
06-JAN-2000; 2000MO-US00376.
11-FEB-2000; 2000MO-US03565.
18-FEB-2000; 2000MO-US04341.
22-FEB-2000; 2000MO-US04414.
24-FEB-2000; 2000MO-US04914.
24-FEB-2000; 2000MO-US05004.
02-MAR-2000; 2000MO-US05841.
15-MAR-2000; 2000MO-US06884.
20-MAR-2000; 2000MO-US07377.

(GENTECH) GENENTECH INC.

Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
Ferrira N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
Giemiaid CJ, Gunney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
Zhang Z;

WPI: 2001-032160/04.
P-PDB: AAB65177.

PRO polynucleotides used to produce polypeptides used to target
bioactive molecules such as toxins, radiolabels or antibodies, to
specific cells, to cause targeted cell death -

Claim 2; Fig 65; 935pp; English.

The present invention describes human secreted and transmembrane PRO
proteins. The PRO proteins have cytostatic activity. The PRO proteins
can be used for targeted delivery of bioactive molecules, such as
toxins, radiolabels or antibodies, that cause cell death. PRO nucleotides
sequences, and their fragments, can be used as hybridisation probes, in
chromosomal and gene mapping, and in the generation of anti-sense RNA
and DNA. They may also be used to produce transgenic animals which are
used to develop and screen therapeutically useful reagents. The PRO
nucleotide and protein sequence can be used for tissue typing and in
treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
AA644270 to AA644470 represent PCR primers and hybridisation probes used
in the isolation of human PRO sequences. AA644087 to AA644469 and
AAB65154 to AAB65300 represent human PRO polynucleotide and protein
sequences given in the exemplification of the present invention.

Sequence 1701 BP; 557 A; 288 C; 381 G; 475 T; 0 other;

Query Match 2.9%; Score 63.2; DB 22; Length 1701;
Best Local Similarity 77.0%; Pred. No. 0.0012;

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

PR 05-SEP-2000; 2000US-0229513

	PR	06-SEP-2000;	2000US-0230437.
	PR	06-SEP-2000;	2000US-0230438.
	PR	08-SEP-2000;	2000US-0231242.
	PR	08-SEP-2000;	2000US-0231243.
	PR	08-SEP-2000;	2000US-0231244.
	PR	08-SEP-2000;	2000US-0231245.
	PR	08-SEP-2000;	2000US-0231411.
	PR	08-SEP-2000;	2000US-0231412.
	PR	08-SEP-2000;	2000US-0231081.
	PR	08-SEP-2000;	2000US-0231082.
	PR	12-SEP-2000;	2000US-0231968.
	PR	14-SEP-2000;	2000US-0232397.
	PR	14-SEP-2000;	2000US-0232398.
	PR	14-SEP-2000;	2000US-0232399.
	PR	14-SEP-2000;	2000US-0232400.
	PR	14-SEP-2000;	2000US-0232401.
	PR	14-SEP-2000;	2000US-0233064.
	PR	14-SEP-2000;	2000US-0233065.
	PR	21-SEP-2000;	2000US-0234223.
	PR	21-SEP-2000;	2000US-0234224.
	PR	25-SEP-2000;	2000US-0234997.
	PR	25-SEP-2000;	2000US-0234998.
	PR	26-SEP-2000;	2000US-0235484.
	PR	27-SEP-2000;	2000US-0235834.
	PR	27-SEP-2000;	2000US-0235836.
	PR	29-SEP-2000;	2000US-0236327.
	PR	29-SEP-2000;	2000US-0236367.
	PR	29-SEP-2000;	2000US-0236368.
	PR	29-SEP-2000;	2000US-0236369.
	PR	29-SEP-2000;	2000US-0236370.
	PT	02-OCT-2000;	2000US-0236802.
	PT	02-OCT-2000;	2000US-0237037.
	XX	02-OCT-2000;	2000US-0237038.
	PS	02-OCT-2000;	2000US-0237039.
	XX	13-OCT-2000;	2000US-0239935.
	PR	13-OCT-2000;	2000US-0239937.
	PR	20-OCT-2000;	2000US-0240960.
	PR	20-OCT-2000;	2000US-0241221.
	PR	20-OCT-2000;	2000US-0241785.
	PR	20-OCT-2000;	2000US-0241786.
	PR	20-OCT-2000;	2000US-0241787.
	PR	20-OCT-2000;	2000US-0241808.
	PR	20-OCT-2000;	2000US-0241809.
	PR	20-OCT-2000;	2000US-0241826.
	PR	01-NOV-2000;	2000US-0244617.
	PR	08-NOV-2000;	2000US-0246474.
	PR	08-NOV-2000;	2000US-0246475.
	PR	08-NOV-2000;	2000US-0246476.
	PR	08-NOV-2000;	2000US-0246477.
	PR	08-NOV-2000;	2000US-0246478.
	PR	08-NOV-2000;	2000US-0246524.
	PR	08-NOV-2000;	2000US-0246525.
	PR	08-NOV-2000;	2000US-0246526.
	PR	08-NOV-2000;	2000US-0246527.
	PR	08-NOV-2000;	2000US-0246528.
	PR	08-NOV-2000;	2000US-0246532.
	PR	08-NOV-2000;	2000US-0246609.
	PR	08-NOV-2000;	2000US-0246610.
	PR	08-NOV-2000;	2000US-0246611.
	PR	08-NOV-2000;	2000US-0246613.
	PR	17-NOV-2000;	2000US-0249207.
	PR	17-NOV-2000;	2000US-0249208.
	PR	17-NOV-2000;	2000US-0249209.
	PR	17-NOV-2000;	2000US-0249210.
	PR	17-NOV-2000;	2000US-0249211.
	PR	17-NOV-2000;	2000US-0249212.
	PR	17-NOV-2000;	2000US-0249213.
	PR	17-NOV-2000;	2000US-0249214.
	PR	17-NOV-2000;	2000US-0249215.
	PR	17-NOV-2000;	2000US-0249216.
	PR	17-NOV-2000;	2000US-0249217.
	PR	06-SEP-2000;	2000US-0230437.
	PR	06-SEP-2000;	2000US-0230438.
	PR	08-SEP-2000;	2000US-0231242.
	PR	08-SEP-2000;	2000US-0231243.
	PR	08-SEP-2000;	2000US-0231244.
	PR	08-SEP-2000;	2000US-0231245.
	PR	08-SEP-2000;	2000US-0231411.
	PR	08-SEP-2000;	2000US-0231412.
	PR	08-SEP-2000;	2000US-0231081.
	PR	08-SEP-2000;	2000US-0231082.
	PR	12-SEP-2000;	2000US-0231968.
	PR	14-SEP-2000;	2000US-0232397.
	PR	14-SEP-2000;	2000US-0232398.
	PR	14-SEP-2000;	2000US-0232399.
	PR	14-SEP-2000;	2000US-0232400.
	PR	14-SEP-2000;	2000US-0232401.
	PR	14-SEP-2000;	2000US-0233064.
	PR	14	

[illegible]

PS Claim 1; Page 2371; 9803pp; English.

XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 CC
 XX

SO Sequence 705 BP; 258 A; 97 C; 99 G; 247 T; 4 other;

Query Match 2.9%; Score 62; DB 22; Length 705;
 Best Local Similarity 62.5%; Pred. No. 0.0018;

Matches 95; Conservative 1; Mismatches 56; Indels 0; Gaps 0;

OY 1993 agataccattattgagccacattggtgcgtacacatttcttatttc 2052
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 519 agtatagaaactataaacaatataatgattttaaacattttacataagtaaa 578
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 2053 cctaatcattagactctcatattctlaaagaatatttcctgttgtaaaaaaaa 2112
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 579 ctgcatcttctgacataacttaacttaaaataagctgcataattttaaaaaaaa 638
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 2113 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2144
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 639 aaaaaaaaaaaaaaaaaaaaaaaaaaaaar 670

Search completed: July 25, 2002, 09:16:35
 Job time: 8225 sec

